

PROGRAMA OFICIAL DE POSTGRADO EN ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

DATOS IDENTIFICATIVOS DE LA MATERIA

Código de la materia: P1061219

Nombre de la materia: TÉCNICAS DE REMUESTREO

Número de créditos ECTS: 5 (Teóricos: 3, Aplicados: 2)

Curso académico: 2014/2015

Profesorado: Ricardo Cao Abad

OBJETIVOS DE LA MATERIA

Se pretende que el alumno adquiera destreza en la identificación de situaciones en las que los métodos de remuestreo son herramientas inferenciales adecuadas para resolver problemas reales. Para ello se tratará de que el alumno conozca el funcionamiento de las principales técnicas de remuestreo, entre las que se destaca el método bootstrap, así como sus aplicaciones en los principales ámbitos de la estadística. Asimismo se persigue que el alumno sea capaz de diseñar e implementar en ordenador planes de remuestreo adecuados para un amplio abanico de situaciones.

CONTENIDOS DE LA MATERIA

1. Motivación del principio Bootstrap. El Bootstrap uniforme. Cálculo de la distribución Bootstrap: distribución exacta y distribución aproximada por Monte Carlo. Ejemplos.
2. Aplicación del Bootstrap a la estimación de la precisión y el sesgo de un estimador. Ejemplos.
3. Motivación del método Jackknife. Estimación Jackknife de la precisión y el sesgo de un estimador. Relación Bootstrap/Jackknife en dicha estimación. Ejemplos. Estudios de simulación.

4. Modificaciones del Bootstrap uniforme: Bootstrap paramétrico, simetrizado, suavizado, ponderado y sesgado. Discusión y ejemplos. Validez de la aproximación Bootstrap. Ejemplos.
5. Aplicación del Bootstrap a la construcción de intervalos de confianza: Métodos percentil, percentil-t, percentil-t simetrizado. Ejemplos. Estudios de simulación.
6. Bootstrap y estimación no paramétrica de la densidad. Aproximación Bootstrap de la distribución del estimador de Parzen-Rosenblatt. El Bootstrap en la selección del parámetro de suavizado.
7. Otras aplicaciones del Bootstrap: El Bootstrap de un modelo de regresión. Wild Bootstrap. El Bootstrap en la estimación del error de predicción. El Bootstrap en poblaciones finitas. Ejemplos. Estudios de simulación.
8. Bootstrap y estimación no paramétrica de la función de regresión. Aproximación Bootstrap de la distribución del estimador de Nadaraya-Watson. Distintos métodos de remuestreo y resultados para ellos.
9. El Bootstrap con datos censurados. Introducción a los datos censurados. Remuestreos Bootstrap en presencia de censura. Relaciones entre ellos.
10. El Bootstrap con datos dependientes. Introducción a las condiciones de dependencia y modelos habituales de datos dependientes. Modelos paramétricos de dependencia. Situaciones de dependencia general: el Bootstrap por bloques, el Bootstrap estacionario y el método del submuestreo.
11. Iteración del principio Bootstrap. Motivación y principales resultados. Aplicaciones del Bootstrap iterado: Corrección del sesgo de un estimador. Corrección del error de cobertura de un intervalo de confianza. Estudios de simulación.

BIBLIOGRAFÍA BÁSICA Y COMPLEMENTARIA

Bibliografía básica

Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and their Application*. Cambridge University Press.

Efron, B. (1979). Bootstrap Methods: Another look at the Jackknife. *Ann. Statist.*, 7, 1-26.

Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall.

Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap*. Springer Verlag.

Bibliografía complementaria

Akritis, M. G. (1986). Bootstrapping the Kaplan--Meier estimator. *J. Amer. Statist. Assoc.* 81, 1032-1038.

Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* 12, 470-482.

Bühlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli* 3, 123-148.

Cao, R. (1990). Órdenes de convergencia para las aproximaciones normal y bootstrap en la estimación no paramétrica de la función de densidad. *Trabajos de Estadística*, vol. 5, 2, 23-32.

Cao, R. (1991). Rate of convergence for the wild bootstrap in nonparametric regression. *Ann. Statist.* 19, 2226-2231.

Cao, R. and Prada-Sánchez, J.M. (1993). Bootstrapping the mean of a symmetric population. *Statistics & Probability Letters* 17, 43-48.

Cao, R. (1993). Bootstrapping the mean integrated squared error. *Jr. Mult. Anal.* 45, 137-160.

Cao, R. (1999). An overview of bootstrap methods for estimating and predicting in time series. *Test*, 8, 95-116.

Cao, R. and González-Manteiga, W. (1993). Bootstrap methods in regression smoothing. *J. Nonparam. Statist.* 2, 379-388.

Efron, B. (1981). Censored data and the bootstrap. *J. Amer. Statist. Assoc.* 76, 312-319.

Efron, B. (1982). The Jackknife, the Bootstrap and other Resampling Plans. *CBMS-NSF. Regional Conference series in applied mathematics.*

Efron, B. (1983). Estimating the error rate of a prediction rule: improvements on cross-validation. *J. Amer. Stat. Assoc.* 78, 316-331.

Efron, B. and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1, 54-77.

Efron, B. (1987). Better Bootstrap confidence intervals (with discussion), *J. Amer. Stat. Assoc.* 82, 171-200.

Efron, B. (1990). More Efficient Bootstrap Computations. *J. Amer. Stat. Assoc.* 85, 79-89.

Freedman, D.A. (1981). Bootstrapping regression models. *Ann. Statist.* 9, 6, 1218-1228.

González-Manteiga, W. y Prada-Sánchez, J.M. (1985). Una aplicación de los métodos de suavización no paramétricos en la técnica Bootstrap. *Proceedings Jornadas Hispano-Lusas de Matemáticas*. Murcia.

García-Jurado, I. González-Manteiga, W., Prada-Sánchez, J.M., Febrero-Bande, M. and Cao, R. (1995). Predicting using Box-Jenkins, nonparametric and bootstrap techniques. *Technometrics* 37, 303-310.

González-Manteiga, W., Prada-Sánchez, J.M. and Romo, J. (1994). The Bootstrap-A Review. *Computational Statistics*, 9, 165-205.

Hall, P. (1986). On the bootstrap and confidence intervals. *Ann. Statist.* 14, 1431-1452.

Hall, P. (1988-a). Theoretical comparison of bootstrap confidence intervals. *Ann. Statist.* 16, 927-953.

Hall, P. (1988-b). Rate of convergence in bootstrap approximations. *Ann. Probab.* 16, 4, 1665-1684.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer Verlag.

Hall, P. and Martin, M.A. (1988). On bootstrap resampling and iteration. *Biometrika* 75, 661-671.

Härdle, W. and Marron, J. S. (1991). Bootstrap simultaneous error bars for nonparametric regression. *Ann. Statist.* 19, 778-796.

Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* 17, 1217-1241.

Lombardía, M.J., González-Manteiga, W. and Prada-Sánchez, J.M. (2003). Bootstrapping the Chambers-Dunstan estimate of a finite population distribution function. *J. Stat. Plan. Infer.*, 116, 367-388.

Mammen, E. (1992). *When does Bootstrap Work?*. Springer Verlag.

Navidi, W. (1989). Edgeworth expansions for bootstrapping regression models. *Ann. Statist.* 17, 4, 1472-1478.

Politis, D.N. and Romano, J.R. (1994). The stationary bootstrap. *J. Amer. Statist. Assoc.* 89, 1303-1313.

Politis, D.N. and Romano, J.R. (1994). Limit theorems for weakly dependent Hilbert space valued random variables with application to the stationary bootstrap. *Statist. Sin.* 4, 461-476.

Politis, D.N., Romano, J.P. and Wolf, M. (1999). *Subsampling*. Springer Verlag.

Prada-Sánchez, J.M. and Otero-Cepeda, X.L. (1989). The use of smooth bootstrap techniques for estimating the error rate of a prediction rule. *Comm. Statist. - Simula.*, 18(3), 1169-1186.

Prada-Sánchez, J.M. and Cotos-Yáñez, T. (1997). A Simulation Study of Iterated and Non-iterated Bootstrap Methods for Bias Reduction and Confidence Interval Estimation. *Comm. Statist. - Simula.*, 26(3), 927-946.

Reid, N. (1981). Estimating the median survival time. *Biometrika* 68, 601-608.

Stine, R.A. (1987). Estimating properties of autoregressive forecasts. *J. Amer. Statist. Assoc.* 82, 1072-1078.

Thombs, L.A. and Schucany, W.R. (1990). Bootstrap prediction intervals for autoregression. *J. Amer. Statist. Assoc.* 85, 486-492.

Wu, C.-F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Statist.* 14, 1261-1350.

COMPETENCIAS GENERALES Y ESPECÍFICAS

Competencias generales:

G1 - Capacidad para iniciar la investigación y para participar en proyectos de investigación que pueden culminar en la elaboración de una tesis doctoral.

G2 - Capacidad de aplicación de algoritmos de resolución de los problemas y manejo del software adecuado.

G3 - Capacidad de trabajo en equipo y de forma autónoma

G4 - Capacidad de formular problemas en términos estadísticos, y de resolverlos utilizando las técnicas adecuadas.

G6 - Capacidad de identificar y resolver problemas

G10 - Capacidad de integrarse en un equipo multidisciplinar para el análisis experimental

G11 - Adquirir destreza para el desarrollo de software

G12 - Capacidad de análisis estadístico crítico de las muestras, los planteamientos y resultados

G13 - Redacción de informes estadísticos con precisión, orden y claridad

G14 - Representar un problema real mediante un modelizado estadístico adecuado.

G15 - Diseñar un plan de observación o recogida de datos que permita abordar el problema de interés

Competencias específicas:

E2 - La adquisición de los conocimientos de estadística e investigación operativa necesarios para la incorporación en equipos multidisciplinarios pertenecientes a diferentes sectores profesionales.

E4 - Conocer las aplicaciones de los modelos de la estadística y la investigación operativa.

- E5 - Conocer algoritmos de resolución de los problemas y manejar el software adecuado.
- E11 - Modelar la dependencia entre una variable respuesta (dependiente) y varias variables explicativas (independientes).
- E12 - Realizar inferencias respecto a los parámetros que aparecen en el modelo.
- E19 - Tratamiento de datos y análisis estadístico de los resultados obtenidos.
- E22 - Capacidad de identificar y resolver problemas que requieran el uso de técnicas del análisis de series de tiempo.
- E27 - Obtener los conocimientos precisos para un análisis crítico y riguroso de los resultados.
- E28 - Complementar el aprendizaje de los aspectos metodológicos con apoyo de software.
- E82 - El estudiante será capaz de comprender la importancia de la Inferencia Estadística como herramienta de obtención de información sobre la población en estudio, a partir del conjunto de datos observados de una muestra representativa de ésta. Para ello deberá reconocer la diferencia entre estadística paramétrica y no paramétrica.
- E84 - Ser capaz de manejar diverso software (en particular R) e interpretar los resultados que proporcionan éstos en los correspondientes estudios prácticos.
- E86 - Soltura en el manejo de la teoría de la probabilidad y las variables aleatorias.
- E78 - Fomentar la sensibilidad hacia los principios del pensamiento científico, favoreciendo las actitudes asociadas al desarrollo de los métodos matemáticos, como: el cuestionamiento de las ideas intuitivas, el análisis crítico de las afirmaciones, la capacidad de análisis y síntesis o la toma de decisiones racionales

METODOLOGÍA DOCENTE: ACTIVIDADES DE APRENDIZAJE Y SU VALORACIÓN EN CRÉDITOS ECTS

Tres quintas partes de la docencia presencial se impartirán mediante exposiciones orales del profesor mientras que el resto corresponderá a prácticas, propuestas por el profesor (se hará uso del paquete estadístico R; por tanto, es necesario que los alumnos dispongan en el aula de un ordenador). El total de ambas actividades tendrá una valoración de 1.5 créditos ECTS. Los 3.5 créditos ECTS restantes corresponderán a estudio personal (2 créditos) y realización de trabajos individuales que el alumno tendrá que elaborar a lo largo del curso (1.5 créditos).

CRITERIOS Y MÉTODOS DE EVALUACIÓN

La evaluación se realizará por medio de prácticas propuestas por el profesor, un trabajo individual a realizar por el alumno, así como un examen escrito. La calificación del examen escrito representará el 40% de la calificación global, las prácticas corresponderán al 20% mientras que el 40% restante corresponderá al trabajo individual, que ha de ser presentado en público por los alumnos.

TIEMPO DE ESTUDIO Y DE TRABAJO PERSONAL QUE DEBE DEDICAR UN ESTUDIANTE PARA SUPERAR LA MATERIA

Docencia presencial: 40 h (24 h de lección magistral y 16 h de prácticas con ordenador).

Estudio y trabajo personal: 85 h

RECOMENDACIONES PARA EL ESTUDIO DE LA MATERIA

Para superar con éxito la materia es aconsejable la asistencia a las clases, siendo fundamental el seguimiento diario del trabajo realizado en el aula. Un prerrequisito necesario es haber cursado al menos una de las materias *Estadística Aplicada* o *Estadística Matemática* del presente master. Es recomendable (aunque no necesario) que el alumno haya cursado la materia optativa de *Simulación Estadística*.

RECURSOS PARA EL APRENDIZAJE

Bibliografía, apuntes y ordenador. Uso del repositorio de material docente del máster.

OBSERVACIONES